A novel efficient mixed formulation for strain-gradient models

DOI: xxx/xxxx

RESEARCH ARTICLE

A novel efficient mixed formulation for strain-gradient models

Stefanos-Aldo Papanicolopulos*[1] | Fahad Gulib[1] | Aikaterini Marinelli[2]

[1]School of Engineering, The University of Edinburgh, Edinburgh, UK
[2]School of Engineering and the Built Environment, Edinburgh Napier University, Edinburgh, UK

Correspondence
*Stefanos-Aldo Papanicolopulos, School of Engineering, The University of Edinburgh, Edinburgh, EH9 3JL, UK. Email: S.Papanicolopulos@ed.ac.uk

Summary

Various finite elements based on mixed formulations have been proposed for the solution of boundary value problems involving strain-gradient models. The relevant literature, however, does not provide details on some important theoretical aspects of these elements. In this work we first present the existing elements within a novel, single mathematical framework, identifying some theoretical issues common to all of them that affect their robustness and numerical efficiency. We then proceed to develop a new family of mixed elements that addresses these issues, while being simpler and computationally cheaper. The behaviour of the new elements is further demonstrated through a numerical example.

KEYWORDS:
Strain-gradient models, Finite element method, mixed formulation, penalty formulation

1 | INTRODUCTION

Starting with seminal work in the 1960s [1,2,3] and continuing up to today [4], strain-gradient theories receive significant interest, being used for example to model the size effect in elastic deformation [5,6,7,8,9] and fracture [10,11,12], and to model shear localisation in plastic deformation [13,14,15,16]. The implementation of such theories in the finite element method is complicated by the presence of strain gradients, that is of second derivatives of the displacements, which in a displacement-only finite-element formulation lead to the requirement for elements with C^1 interpolation. Appropriate C^1 elements have been successfully presented and used [14,7,8], which have the advantage that all degrees of freedom contribute to the interpolation of the displacement field [17].

To overcome the restrictions in element choice imposed by the C^1 requirement, alternative element formulations have been proposed, especially using mixed formulations based on Lagrange multipliers [5,10,15,16,18] or penalty methods [16,19,7]. Elements based on mixed formulations have also been developed for couple stress theories [20,21], where similar issues exist. These elements achieve simpler interpolation of the displacement field, generally using well-known quadratic interpolations, by introducing an additional "displacement gradient" field. The relevant literature, however, does not provide details on some important theoretical aspects of these elements and does not allow for an easy comparative study of their similarities and differences.

In Section 2 we introduce a new, general formulation of mixed formulations using Lagrange multipliers. In doing so, we also clarify some aspects of the method regarding the exact or approximative nature of the quantities and equations involved. Section 3 shows how the Lagrange multiplier formulation can be easily converted to a penalty formulation. A different penalty formulation is also derived, which includes the formulation obtained through physical arguments by Zervos [19]. In Section 4 we use our newly presented formulation to revisit the elements proposed in the literature [5,10,15,18], allowing for easy identification of their common characteristics and their differences. This leads to the identification of two theoretical issues with existing elements, related to the dependence on the rotation field. This issue is overcome in Section 5, by developing a new mixed formulation for strain-gradient models that leads to simpler, more robust and more efficient elements in comparison to the existing ones. The improved performance of the new elements is also demonstrated in Section 6 through specific numerical simulations.

2 | GENERAL MIXED FORMULATION FOR STRAIN-GRADIENT THEORIES

This section presents a general, abstract mixed formulation for strain-gradient theories using Lagrange multipliers, which can describe all elements of this type presented in the literature.

In the quasi-static, small-strain, Form I strain gradient theory[2,3], the weak form of the equilibrium equation can be written as

$$\int_V (\sigma_{ij}\delta\epsilon_{ij} + \tilde{\mu}_{ijk}\delta\tilde{\kappa}_{ijk})\,\mathrm{d}V = \int_V F_k\delta u_k\,\mathrm{d}V + \int_S P_k\delta u_k\,\mathrm{d}S + \int_S R_k\hat{n}_i(\delta u_k)_{,i}\,\mathrm{d}S + \oint_C E_k\delta u_k\,\mathrm{d}C \tag{1}$$

The kinematic quantities in equation (1) are the displacement u_k, the strain $\epsilon_{ij} \equiv (u_{i,j} + u_{j,i})/2$ and the second gradient of the displacement $\tilde{\kappa}_{ijk} \equiv u_{k,ij}$. The static quantities are the stress σ_{ij} and the double stress $\tilde{\mu}_{ijk}$, which are work conjugate to ϵ_{ij} and $\tilde{\kappa}_{ijk}$ respectively. The external actions are the body force F_k, the surface traction P_k, the surface double traction R_k and the edge traction E_k. Finally, V is the domain occupied by the body under consideration, S is the boundary of V, C are the edges of S, and $\hat{n}_i$ is the outward unit normal to S.

Like its classical counterpart, this weak form has the general matrix form

$$\int_V \delta\boldsymbol{\epsilon}_g^T\boldsymbol{\sigma}_g\,\mathrm{d}V = \boldsymbol{\mathcal{F}}\delta\boldsymbol{u} \tag{2}$$

where we have introduced the *generalised strain vector* $\boldsymbol{\epsilon}_g$ and the *generalised stress vector* $\boldsymbol{\sigma}_g$, and where $\boldsymbol{\mathcal{F}}$ is a linear operator (in row vector form). The vector $\boldsymbol{\epsilon}_g$ only depends on the displacement $\boldsymbol{u}$; for a second-gradient theory, it is a function of the first and second derivatives of the displacement $\boldsymbol{u}$. The same matrix form applies to Form II and Form III of the strain gradient theory[2,3], so the rest of the presentation in this section is valid for all forms.

In the most general case, individual elements of $\boldsymbol{\epsilon}_g$ could depend on both first and second derivatives. We choose to consider separately the first and the second derivatives of $\boldsymbol{u}$, as this is always the case in practice, so we can write

$$\boldsymbol{\epsilon}_g = \begin{bmatrix}\boldsymbol{\epsilon}\\ \boldsymbol{\kappa}\end{bmatrix} = \begin{bmatrix}\boldsymbol{\mathcal{L}}_\epsilon\\ \boldsymbol{\mathcal{L}}_\kappa\boldsymbol{\mathcal{L}}_\upsilon\end{bmatrix}\boldsymbol{u} \tag{3}$$

where $\boldsymbol{\mathcal{L}}_\epsilon$, $\boldsymbol{\mathcal{L}}_\kappa$ and $\boldsymbol{\mathcal{L}}_\upsilon$ are linear operator matrices involving only first spatial derivatives. The generalised stresses are similarly split as

$$\boldsymbol{\sigma}_g = \begin{bmatrix}\boldsymbol{\sigma}\\ \boldsymbol{\mu}\end{bmatrix} \tag{4}$$

Setting

$$\boldsymbol{\upsilon} = \boldsymbol{\mathcal{L}}_\upsilon\boldsymbol{u} \tag{5}$$

and introducing the *extended displacement* vector

$$\boldsymbol{u}_e = \begin{bmatrix}\boldsymbol{u}\\ \boldsymbol{\upsilon}\end{bmatrix} \tag{6}$$

equations (3) and (5) can be written as

$$\boldsymbol{\epsilon}_g = \begin{bmatrix}\boldsymbol{\mathcal{L}}_\epsilon & \cdot\\ \cdot & \boldsymbol{\mathcal{L}}_\kappa\end{bmatrix}\begin{bmatrix}\boldsymbol{u}\\ \boldsymbol{\upsilon}\end{bmatrix} = \boldsymbol{\mathcal{L}}_g\boldsymbol{u}_e \tag{7}$$

$$\boldsymbol{0} = \begin{bmatrix}-\boldsymbol{\mathcal{L}}_\upsilon & \boldsymbol{I}\end{bmatrix}\begin{bmatrix}\boldsymbol{u}\\ \boldsymbol{\upsilon}\end{bmatrix} = \boldsymbol{\mathcal{L}}_c\boldsymbol{u}_e \tag{8}$$

where $\boldsymbol{I}$ is the identity operator (or identity matrix) and a dot $(\cdot)$ is used to indicate a zero submatrix of appropriate dimensions. As seen by comparing equations (1) and (2), the operator $\boldsymbol{\mathcal{F}}$ includes first derivatives, so that we can split it to write

$$\boldsymbol{\mathcal{F}}\delta\boldsymbol{u} = \begin{bmatrix}\boldsymbol{\mathcal{F}}_u & \boldsymbol{\mathcal{F}}_\upsilon\end{bmatrix}\begin{bmatrix}\delta\boldsymbol{u}\\ \delta\boldsymbol{\upsilon}\end{bmatrix} = \boldsymbol{\mathcal{F}}_e\delta\boldsymbol{u}_e \tag{9}$$

Introducing an appropriately-sized vector of Lagrange multipliers $\boldsymbol{\lambda}$, the variation of the product of equation (8) and $\boldsymbol{\lambda}$ yields

$$\delta(\boldsymbol{\lambda}^T\boldsymbol{\mathcal{L}}_c\boldsymbol{u}_e) = (\delta\boldsymbol{\lambda}^T)\boldsymbol{\mathcal{L}}_c\boldsymbol{u}_e + \delta(\boldsymbol{\mathcal{L}}_c\boldsymbol{u}_e)^T\boldsymbol{\lambda} = 0 \tag{10}$$

We can therefore add this quantity to $\delta\boldsymbol{\epsilon}_g^T\boldsymbol{\sigma}_g$ without changing its value, so that

$$\delta\boldsymbol{\epsilon}_g^T\boldsymbol{\sigma}_g = \delta\boldsymbol{\epsilon}_g^T\boldsymbol{\sigma}_g + (\delta\boldsymbol{\lambda}^T)\boldsymbol{\mathcal{L}}_c\boldsymbol{u}_e + \delta(\boldsymbol{\mathcal{L}}_c\boldsymbol{u}_e)^T\boldsymbol{\lambda} = \delta\boldsymbol{\epsilon}_o^T\boldsymbol{\sigma}_o \tag{11}$$

thus introducing the *overall strain*, *overall stress* and *overall displacement* vectors

$$\boldsymbol{\epsilon}_o \equiv \begin{bmatrix} \boldsymbol{\epsilon}_g \\ \boldsymbol{\mathcal{L}}_c \boldsymbol{u}_e \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mathcal{L}}_g \boldsymbol{u}_e \\ \boldsymbol{\mathcal{L}}_c \boldsymbol{u}_e \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mathcal{L}}_g & \cdot \\ \boldsymbol{\mathcal{L}}_c & \cdot \\ \cdot & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_e \\ \boldsymbol{\lambda} \end{bmatrix} = \boldsymbol{\mathcal{L}}_o \boldsymbol{u}_o, \ \boldsymbol{\sigma}_o \equiv \begin{bmatrix} \boldsymbol{\sigma}_g \\ \boldsymbol{\lambda} \\ \boldsymbol{\mathcal{L}}_c \boldsymbol{u}_e \end{bmatrix}, \ \boldsymbol{u}_o \equiv \begin{bmatrix} \boldsymbol{u}_e \\ \boldsymbol{\lambda} \end{bmatrix} \tag{12}$$

The weak form of the equilibrium equations can therefore be written as

$$\int_V \delta \boldsymbol{\epsilon}_o^T \boldsymbol{\sigma}_o \, \mathrm{d}V = \boldsymbol{\mathcal{F}}_e \delta \boldsymbol{u}_e = \begin{bmatrix} \boldsymbol{\mathcal{F}}_e & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \delta \boldsymbol{u}_e \\ \delta \boldsymbol{\lambda} \end{bmatrix} = \boldsymbol{\mathcal{F}}_o \delta \boldsymbol{u}_o \tag{13}$$

It is important to realise that this form is equivalent to the original equilibrium equations. Contrary to what is stated in the literature[5,15], the introduction of the Lagrange multipliers enforces the relation (5) between $\boldsymbol{v}$ and $\boldsymbol{u}$ in an exact way, and not only in a weak, volume-averaged sense. This also means that there is no need to introduce the concept of *relaxed*[5] displacement gradient and strain; $\boldsymbol{v}$ actually contains the (true) displacement derivatives.

Following the standard finite element procedure, we now introduce the approximation of the overall displacement field $\boldsymbol{u}_o$ through the discretisation

$$\boldsymbol{u}_o \approx \bar{\boldsymbol{u}}_o = \boldsymbol{N}_o \boldsymbol{u}_o^N \tag{14}$$

where $\boldsymbol{N}_o$ is a matrix of shape functions (depending on the spatial coordinates), $\boldsymbol{u}_o^N$ is a vector of discrete nodal values, that is of degrees of freedom (dofs), and we use a bar $(\bar{\cdot})$ to indicate approximated quantities. The overall strain is approximated as

$$\boldsymbol{\epsilon}_o \approx \bar{\boldsymbol{\epsilon}}_o = \boldsymbol{\mathcal{L}}_o \bar{\boldsymbol{u}}_o = \boldsymbol{\mathcal{L}}_o (\boldsymbol{N}_o \boldsymbol{u}_o^N) = (\boldsymbol{\mathcal{L}}_o \boldsymbol{N}_o) \boldsymbol{u}_o^N = \boldsymbol{B}_o \boldsymbol{u}_o^N \tag{15}$$

The discretised version of the weak form of the equilibrium equations (13) is therefore

$$(\delta \boldsymbol{u}_o^N)^T \int_V \boldsymbol{B}_o^T \boldsymbol{\sigma}_o \, \mathrm{d}V = (\delta \boldsymbol{u}_o^N)^T (\boldsymbol{\mathcal{F}}_o \boldsymbol{N}_o)^T \tag{16}$$

which, as it must hold for arbitrary $\delta \boldsymbol{u}_o^N$, yields the system of equations

$$\boldsymbol{r}_o = (\boldsymbol{\mathcal{F}}_o \boldsymbol{N}_o)^T - \int_V \boldsymbol{B}_o^T \boldsymbol{\sigma}_o \, \mathrm{d}V = \boldsymbol{0} \tag{17}$$

To proceed further, we choose to interpolate separately the extended displacements and the Lagrange multipliers, to obtain

$$\bar{\boldsymbol{u}}_o = \begin{bmatrix} \bar{\boldsymbol{u}}_e \\ \bar{\lambda} \end{bmatrix} = \begin{bmatrix} \boldsymbol{N}_e & \cdot \\ \cdot & \boldsymbol{N}_\lambda \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_e^N \\ \lambda^N \end{bmatrix} \tag{18}$$

where again $\boldsymbol{N}_e$ and $\boldsymbol{N}_\lambda$ are matrices of shape functions, while $\boldsymbol{u}_e^N$ and $\boldsymbol{\lambda}^N$ are vectors of discrete nodal values. We then compute

$$\boldsymbol{B}_o = \boldsymbol{\mathcal{L}}_o \boldsymbol{N}_o = \begin{bmatrix} \boldsymbol{\mathcal{L}}_g & \cdot \\ \boldsymbol{\mathcal{L}}_c & \cdot \\ \cdot & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{N}_e & \cdot \\ \cdot & \boldsymbol{N}_\lambda \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mathcal{L}}_g \boldsymbol{N}_e & \cdot \\ \boldsymbol{\mathcal{L}}_c \boldsymbol{N}_e & \cdot \\ \cdot & \boldsymbol{N}_\lambda \end{bmatrix} = \begin{bmatrix} \boldsymbol{B}_g & \cdot \\ \boldsymbol{B}_c & \cdot \\ \cdot & \boldsymbol{N}_\lambda \end{bmatrix} \tag{19}$$

so that, with simple substitutions and calculations, the residual $\boldsymbol{r}_o$ can be written as

$$\boldsymbol{r}_o = \begin{bmatrix} (\boldsymbol{\mathcal{F}}_e \boldsymbol{N}_e)^T \\ \cdot \end{bmatrix} - \int_V \begin{bmatrix} \boldsymbol{B}_g^T \boldsymbol{\sigma}_g \\ \cdot \end{bmatrix} \mathrm{d}V - \left(\int_V \begin{bmatrix} \cdot & \boldsymbol{B}_c^T \boldsymbol{N}_\lambda \\ \boldsymbol{N}_\lambda^T \boldsymbol{B}_c & \cdot \end{bmatrix} \mathrm{d}V \right) \begin{bmatrix} \boldsymbol{u}_e^N \\ \lambda^N \end{bmatrix} = \begin{bmatrix} \boldsymbol{r}_g \\ \cdot \end{bmatrix} - \begin{bmatrix} \cdot & \boldsymbol{K}_c \\ \boldsymbol{K}_c^T & \cdot \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_e^N \\ \lambda^N \end{bmatrix} \tag{20}$$

Within an iterative solution of the, generally non-linear, system (17), we need to compute the stiffness matrix (Jacobian)

$$\boldsymbol{K} = -\frac{\partial \boldsymbol{r}_o}{\partial \boldsymbol{u}_o^N} = \int_V \begin{bmatrix} \boldsymbol{B}_g^T \boldsymbol{D}_g \boldsymbol{B}_g & \boldsymbol{B}_c^T \boldsymbol{N}_\lambda \\ \boldsymbol{N}_\lambda^T \boldsymbol{B}_c & \cdot \end{bmatrix} \mathrm{d}V = \begin{bmatrix} \boldsymbol{K}_e & \boldsymbol{K}_c \\ \boldsymbol{K}_c^T & \cdot \end{bmatrix} \tag{21}$$

where the linearisation moduli $\boldsymbol{D}_g$ are defined as

$$\boldsymbol{D}_g = \frac{\partial \boldsymbol{\sigma}_g}{\partial \boldsymbol{\epsilon}_g} \tag{22}$$

so that we obtain the iterative update by solving the linear system

$$\boldsymbol{K} d\boldsymbol{u}_o = \boldsymbol{r}_o \quad \Rightarrow \quad \begin{bmatrix} \boldsymbol{K}_e & \boldsymbol{K}_c \\ \boldsymbol{K}_c^T & \cdot \end{bmatrix} \begin{bmatrix} d\boldsymbol{u}_e^N \\ d\lambda^N \end{bmatrix} = \begin{bmatrix} \boldsymbol{r}_g \\ \cdot \end{bmatrix} - \begin{bmatrix} \cdot & \boldsymbol{K}_c \\ \boldsymbol{K}_c^T & \cdot \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_e^N \\ \lambda^N \end{bmatrix} \tag{23}$$

Since $\boldsymbol{B}_g$ and $\boldsymbol{B}_c$ include only first derivatives of the shape functions $\boldsymbol{N}_e$, this interpolation only needs to be C^0 continuous. This is the main advantage of the proposed mixed elements, allowing the use of well-known C^0 shape functions, while a displacement-only formulation would require the use of a C^1 interpolation. There are no derivatives of $\boldsymbol{N}_\lambda$ involved, so the interpolation of the Lagrange multipliers can be discontinuous. On the other hand, the proposed approach has two disadvantages: the first one is that we introduce degrees of freedom for the Lagrange multipliers and those have corresponding zeroes on the diagonal (the system is no longer positive definite), and the second one is that we need to separately interpolate a second field ($\boldsymbol{v}$) which in its exact form does not provide more information than the $\boldsymbol{u}$ field.

A further difficulty of the mixed formulation is that the theoretical analysis of its stability with respect to mesh refinement requires proving that the Babuška-Brezzi (inf-sup) conditions are satisfied[22,23]. This is not straightforward and indeed, to our knowledge, has not been done for any of the existing mixed finite elements presented in this paper. It is worth noting, however, that Zienkiewicz and Taylor[24] state that "…with certain restrictions on regularity of the problem dealt with, the satisfaction of the patch test is equivalent to the satisfaction of the more mathematical Babuška-Brezzi criteria".

The procedure described in this section applies independently of the space dimensions or of the element geometry. Additionally, equations (20) and (21) have the same form both for the entire domain and for an individual finite element.

The general formulation described in this section does not require specifying the form of $\boldsymbol{N}_e$. In practice, $\boldsymbol{u}$ and $\boldsymbol{v}$ will be interpolated separately so that

$$\boldsymbol{N}_e = \begin{bmatrix} \boldsymbol{N}_u & \cdot \\ \cdot & \boldsymbol{N}_v \end{bmatrix} \qquad \Rightarrow \qquad \bar{\boldsymbol{u}}_e = \begin{bmatrix} \bar{\boldsymbol{u}} \\ \bar{\boldsymbol{v}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{N}_u \boldsymbol{u}^N \\ \boldsymbol{N}_v \boldsymbol{v}^N \end{bmatrix} \tag{24}$$

As mentioned above, the elements of $\boldsymbol{v}$ are the derivatives of the displacements. The nodal values $\boldsymbol{v}^N$, however, are not exactly equal to the derivatives of $\boldsymbol{u}$ calculated at the nodes. While this may at first seem counter-intuitive, and an indication that it is necessary to introduce relaxed versions of the displacement derivatives, it must be noted that $\boldsymbol{u}^N$ are not the values of $\boldsymbol{u}$ calculated at the nodes either. On the contrary, $\boldsymbol{u}^N$ are the values of the approximated displacements $\bar{\boldsymbol{u}}$ evaluated at the nodes, and similarly $\boldsymbol{v}^N$ are the values of the approximated displacement derivatives $\bar{\boldsymbol{v}}$ evaluated at the nodes. It would therefore be more appropriate to use the notation $\bar{\boldsymbol{u}}^N$ and $\bar{\boldsymbol{v}}^N$ for the nodal values, but we choose not to do so to keep the notation cleaner.

3 | PENALTY METHOD FORMULATION

To avoid the zero diagonal block in equation (23), we modify equation (20) as follows

$$\boldsymbol{r}_o \approx \boldsymbol{r}_{o|p} = \begin{bmatrix} \boldsymbol{r}_g \\ \cdot \end{bmatrix} - \begin{bmatrix} \cdot & \boldsymbol{K}_c \\ \boldsymbol{K}_c^T & -(\gamma \boldsymbol{G})^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_e^N \\ \lambda^N \end{bmatrix} \tag{25}$$

where $\boldsymbol{G}$ is an invertible symmetric matrix and γ is a large number, so that as $\gamma \to \infty$ we get $\boldsymbol{r}_{o|p} \to \boldsymbol{r}_o$. To solve the system $\boldsymbol{r}_{o|p} = \boldsymbol{0}$ we can directly solve for λ^N to obtain

$$\lambda^N = \gamma \boldsymbol{G} \boldsymbol{K}_c^T \boldsymbol{u}_e^N \tag{26}$$

so the system to solve becomes $\boldsymbol{r}_e = \boldsymbol{0}$, where

$$\boldsymbol{r}_e = \boldsymbol{r}_g - \gamma \boldsymbol{K}_c \boldsymbol{G} \boldsymbol{K}_c^T \boldsymbol{u}_e^N \tag{27}$$

Within an iterative procedure, we get the finite element equation

$$(\boldsymbol{K}_e + \gamma \boldsymbol{K}_c \boldsymbol{G} \boldsymbol{K}_c^T) d\boldsymbol{u}_e^N = \boldsymbol{r}_g - \gamma \boldsymbol{K}_c \boldsymbol{G} \boldsymbol{K}_c^T \boldsymbol{u}_e^N \tag{28}$$

The penalty method formulation is much simpler to implement than the one with Lagrange multipliers, since the stiffness matrix in equation (28) is positive definite and there are no nodal values for the Lagrange multipliers. It is necessary, however, to determine an appropriate value of the penalty parameter γ and the matrix $\boldsymbol{G}$ to minimise the additional numerical error.

An alternative penalty formulation can be obtained by introducing directly the penalty term, so that the weak form of the equilibrium equation becomes

$$\int_V \left(\delta \boldsymbol{\epsilon}_g^T \boldsymbol{\sigma}_g + \gamma \delta (\boldsymbol{\mathcal{L}}_c \boldsymbol{u}_e)^T \boldsymbol{G} (\boldsymbol{\mathcal{L}}_c \boldsymbol{u}_e) \right) \mathrm{d}V = \boldsymbol{\mathcal{F}}_e \delta \boldsymbol{u}_e \tag{29}$$

which is obviously equivalent to the original weak form, due to equation (8). Using the approximation $\boldsymbol{u}_e \approx \boldsymbol{N}_e \boldsymbol{u}_e^N$ we obtain after some calculations

$$(\delta \boldsymbol{u}_e^N)^T \int_V \left(\boldsymbol{B}_g^T \boldsymbol{\sigma}_g + \gamma \boldsymbol{B}_c^T \boldsymbol{G} \boldsymbol{B}_c \boldsymbol{u}_e^N \right) \mathrm{d}V = (\delta \boldsymbol{u}_e^N)^T (\boldsymbol{\mathcal{F}}_e \boldsymbol{N}_e)^T \tag{30}$$

Since $\delta \boldsymbol{u}_e^N$ is arbitrary, the system to solve becomes $\boldsymbol{r}'_e = \boldsymbol{0}$, where

$$\boldsymbol{r}'_e = (\boldsymbol{\mathcal{F}}_e \boldsymbol{N}_e)^T - \int_V \boldsymbol{B}_g^T \boldsymbol{\sigma}_g \, \mathrm{d}V - \gamma \left(\int_V \boldsymbol{B}_c^T \boldsymbol{G} \boldsymbol{B}_c \, \mathrm{d}V \right) \boldsymbol{u}_e^N \tag{31}$$

4 | REVISITING EXISTING MIXED ELEMENTS

4.1 | Description of existing elements using the proposed framework

The general procedure outlined above can be used to derive the different mixed elements already presented in the literature. As an example, we see here how to derive the elements presented by Shu *et al.*[5] (ignoring the incompressible case).

These are all two-dimensional plane-strain elements, formulated using Form I of strain-gradient theory (so that $\boldsymbol{\kappa}$ contains the second derivatives of the displacements), and with the elements of $\boldsymbol{v}$ being the components of the displacement gradient. Using the formulation introduced in Section 2, the relevant quantities can be written as

$$\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \ \boldsymbol{\sigma} = \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{12} \end{bmatrix}, \ \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_{11} \\ \epsilon_{22} \\ 2\epsilon_{12} \end{bmatrix}, \ \boldsymbol{\mu} = \begin{bmatrix} \tilde{\mu}_{111} \\ \tilde{\mu}_{221} \\ \tilde{\mu}_{211} \\ \tilde{\mu}_{112} \\ \tilde{\mu}_{222} \\ \tilde{\mu}_{122} \end{bmatrix}, \ \boldsymbol{\kappa} = \begin{bmatrix} \tilde{\kappa}_{111} \\ \tilde{\kappa}_{221} \\ 2\tilde{\kappa}_{211} \\ \tilde{\kappa}_{112} \\ \tilde{\kappa}_{222} \\ 2\tilde{\kappa}_{122} \end{bmatrix} = \begin{bmatrix} u_{1,11} \\ u_{1,22} \\ 2u_{1,12} \\ u_{2,11} \\ u_{2,22} \\ 2u_{2,12} \end{bmatrix}, \ \boldsymbol{v} = \begin{bmatrix} v_{11} \\ v_{12} \\ v_{21} \\ v_{22} \end{bmatrix} = \begin{bmatrix} u_{1,1} \\ u_{1,2} \\ u_{2,1} \\ u_{2,2} \end{bmatrix}, \ \boldsymbol{\lambda} = \begin{bmatrix} \lambda_{11} \\ \lambda_{12} \\ \lambda_{21} \\ \lambda_{22} \end{bmatrix} \tag{32}$$

The basic differential operators are

$$\boldsymbol{\mathcal{L}}_\epsilon = \begin{bmatrix} \partial_1 & \cdot \\ \cdot & \partial_2 \\ \partial_2 & \partial_1 \end{bmatrix}, \ \boldsymbol{\mathcal{L}}_v = \begin{bmatrix} \partial_1 & \cdot \\ \partial_2 & \cdot \\ \cdot & \partial_1 \\ \cdot & \partial_2 \end{bmatrix}, \ \boldsymbol{\mathcal{L}}_\kappa = \begin{bmatrix} \partial_1 & \cdot & \cdot & \cdot \\ \cdot & \partial_2 & \cdot & \cdot \\ \partial_2 & \partial_1 & \cdot & \cdot \\ \cdot & \cdot & \partial_1 & \cdot \\ \cdot & \cdot & \cdot & \partial_2 \\ \cdot & \cdot & \partial_2 & \partial_1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mathcal{L}}_\epsilon & \cdot \\ \cdot & \boldsymbol{\mathcal{L}}_\epsilon \end{bmatrix} \tag{33}$$

where ∂_1 and ∂_2 are the spatial derivatives with respect to the global coordinates x_1 and x_2 respectively.

The nodal quantities are

$$\boldsymbol{u}^N = \begin{bmatrix} \boldsymbol{u}_1^N \\ \boldsymbol{u}_2^N \end{bmatrix}, \ \boldsymbol{v}^N = \begin{bmatrix} \boldsymbol{v}_{11}^N \\ \boldsymbol{v}_{12}^N \\ \boldsymbol{v}_{21}^N \\ \boldsymbol{v}_{22}^N \end{bmatrix}, \ \boldsymbol{\lambda}^N = \begin{bmatrix} \boldsymbol{\lambda}_{11}^N \\ \boldsymbol{\lambda}_{12}^N \\ \boldsymbol{\lambda}_{21}^N \\ \boldsymbol{\lambda}_{22}^N \end{bmatrix} \tag{34}$$

and the shape function matrices are

$$\boldsymbol{N}_u = \begin{bmatrix} \boldsymbol{n}_u & \cdot \\ \cdot & \boldsymbol{n}_u \end{bmatrix}, \ \boldsymbol{N}_v = \begin{bmatrix} \boldsymbol{n}_v & \cdot & \cdot & \cdot \\ \cdot & \boldsymbol{n}_v & \cdot & \cdot \\ \cdot & \cdot & \boldsymbol{n}_v & \cdot \\ \cdot & \cdot & \cdot & \boldsymbol{n}_v \end{bmatrix}, \ \boldsymbol{N}_\lambda = \begin{bmatrix} \boldsymbol{n}_\lambda & \cdot & \cdot & \cdot \\ \cdot & \boldsymbol{n}_\lambda & \cdot & \cdot \\ \cdot & \cdot & \boldsymbol{n}_\lambda & \cdot \\ \cdot & \cdot & \cdot & \boldsymbol{n}_\lambda \end{bmatrix} \tag{35}$$

where $\boldsymbol{n}_u$, $\boldsymbol{n}_v$, $\boldsymbol{n}_\lambda$ are row vectors of shape functions.

Shu *et al.*[5] proposed six different elements, of which the QU34L4 element (see figure 1) performs best. This is a nine-noded isoparametric quadrilateral, that uses a bi-quadratic Lagrangian interpolation for the displacements $\boldsymbol{u}$, a bi-linear interpolation for the displacement gradients $\boldsymbol{v}$ and a constant interpolation for the Lagrange multipliers $\boldsymbol{\lambda}$. Therefore the column vectors $\boldsymbol{u}_{(i)}^N$, $\boldsymbol{v}_{(ij)}^N$ and $\boldsymbol{\lambda}_{(ij)}^N$ have nine, four and one elements respectively, so that $\boldsymbol{u}_e^N$ has 34 elements and $\boldsymbol{\lambda}^N$ has 4 elements, hence the name QU34L4. The element is isoparametric in the sense that the displacement $\boldsymbol{u}$ and the coordinates $\boldsymbol{x}$ are interpolated in the same way, but obviously $\boldsymbol{v}$ and $\boldsymbol{\lambda}$ are interpolated differently. Since we use a constant interpolation for $\boldsymbol{\lambda}$, the matrix $\boldsymbol{N}_\lambda$ is a 4×4 unit matrix. A three-dimensional version of QU34L4, named BR153L9, has also been developed[18].

Matsushima *et al.*[15] have presented a large-strain formulation for plane-strain strain-gradient theories. In the small-strain case, the element they present (see figure 1) is very similar to QU34L4, but using an eight-node serendipity interpolation for the

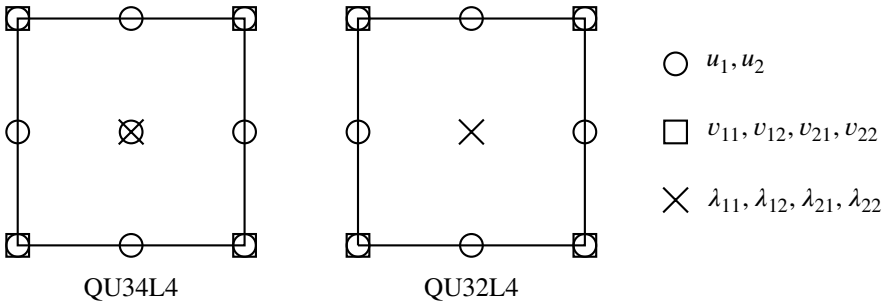

FIGURE 1 Geometry and degrees of freedom for the existing QU34L4 and QU32L4 elements.

displacements. We use here the name QU32L4 for this element, as it has 32 degrees of freedom for $\boldsymbol{u}_e$. The QU34L4 element uses a 3×3 Gauss integration for computing $\boldsymbol{r}_g$ and $\boldsymbol{K}_e$, and a 2×2 Gauss integration for computing $\boldsymbol{K}_c$, while QU32L4 uses a 2×2 Gauss integration for all terms.

Kouznetsova *et al.*[16] present a large-strain formulation, which in the small-strain case yields the same element QU32L4. They also present a penalty formulation, of the form shown in equation (25), with

$$\boldsymbol{G}^{-1} = -\int_V \boldsymbol{N}_\lambda^T \boldsymbol{N}_\lambda \, \mathrm{d}V \tag{36}$$

For elements with non-constant interpolation of $\boldsymbol{\lambda}$, it is necessary to check that the right-hand side of (36) is indeed non-singular.

Zervos[19] introduced a family of two-dimensional and three-dimensional elements for elasticity with microstructure[2] and showed, using physical arguments, that through appropriate choice of material parameters they can be used as a penalty formulation for strain-gradient elasticity. This formulation can be shown to be a special case of the alternative formulation given in equation (31), with the more general formulation presented here having the advantage of being based on mathematical arguments and thus not requiring choice of a specific constitutive model.

The QU34L4 and QU32L4 elements share the general advantages and disadvantages of mixed strain-gradient elements described in Section 2. They have however two additional theoretical drawbacks: spurious zero-element modes for single elements, and dependence on the discretised rotation field.

4.2 | Zero-energy modes

Considering for simplicity a square element aligned with the Cartesian axes, it is easy to see that both elements will have the expected three zero-energy modes (two rigid body translations and one rigid body rotation) plus two spurious zero-energy modes that correspond to the following values of the nodal quantities:

$$\boldsymbol{u}^N = \boldsymbol{0},\ \boldsymbol{v}^N = [1, 1, -1, -1, -1, 1, 1, -1, 0, 0, 0, 0, 0, 0, 0, 0]^T,\ \boldsymbol{\lambda}^N = \boldsymbol{0} \tag{37a}$$

$$\boldsymbol{u}^N = \boldsymbol{0},\ \boldsymbol{v}^N = [0, 0, 0, 0, 0, 0, 0, 0, 1, 1, -1, -1, -1, 1, 1, -1]^T,\ \boldsymbol{\lambda}^N = \boldsymbol{0} \tag{37b}$$

In both cases, these zero-energy modes are not compatible across adjacent elements, so that any multi-element mesh will result in no spurious zero-energy modes.

4.3 | Dependence on the rotation field

The displacement gradient $\boldsymbol{v}$ can be expressed in terms of the strain and rotation fields as

$$\boldsymbol{v} = \begin{bmatrix} v_{11} \\ v_{12} \\ v_{21} \\ v_{22} \end{bmatrix} = \begin{bmatrix} u_{1,1} \\ u_{1,2} \\ u_{2,1} \\ u_{2,2} \end{bmatrix} = \begin{bmatrix} \epsilon_{11} \\ \epsilon_{12} + \omega_{12} \\ \epsilon_{12} - \omega_{12} \\ \epsilon_{22} \end{bmatrix} \tag{38}$$

Therefore the second gradients $\boldsymbol{\kappa}$ are

$$\boldsymbol{\kappa} = \boldsymbol{\mathcal{L}}_\kappa \boldsymbol{v} = \begin{bmatrix} \partial_1 & \cdot & \cdot & \cdot \\ \cdot & \partial_2 & \cdot & \cdot \\ \partial_2 & \partial_1 & \cdot & \cdot \\ \cdot & \cdot & \partial_1 & \cdot \\ \cdot & \cdot & \cdot & \partial_2 \\ \cdot & \cdot & \partial_2 & \partial_1 \end{bmatrix} \begin{bmatrix} \epsilon_{11} \\ \epsilon_{12} + \omega_{12} \\ \epsilon_{12} - \omega_{12} \\ \epsilon_{22} \end{bmatrix} = \begin{bmatrix} \epsilon_{11,1} \\ \epsilon_{12,2} + \omega_{12,2} \\ \epsilon_{11,2} + \epsilon_{12,1} + \omega_{12,1} \\ \epsilon_{12,1} - \omega_{12,1} \\ \epsilon_{22,2} \\ \epsilon_{22,1} + \epsilon_{12,2} - \omega_{12,2} \end{bmatrix} \tag{39}$$

The first gradients of $\boldsymbol{v}$, that is the second gradients of the displacement, are not independent, since

$$v_{11,2} = v_{12,1},\ v_{22,1} = v_{21,2} \tag{40}$$

or, in terms of strains and rotations,

$$\omega_{12,1} = \epsilon_{11,2} - \epsilon_{12,1},\ \omega_{12,2} = \epsilon_{12,2} - \epsilon_{22,1} \tag{41}$$

Substituting (41) into (39) we obtain

$$\boldsymbol{\kappa} = \begin{bmatrix} \epsilon_{11,1} \\ 2\epsilon_{12,2} - \epsilon_{22,1} \\ 2\epsilon_{11,2} \\ 2\epsilon_{12,1} - \epsilon_{11,2} \\ \epsilon_{22,2} \\ 2\epsilon_{22,1} \end{bmatrix} = \begin{bmatrix} \partial_1 & \cdot & \cdot & \cdot \\ \cdot & \partial_2 & \partial_2 & -\partial_1 \\ 2\partial_2 & \cdot & \cdot & \cdot \\ -\partial_2 & \partial_1 & \partial_1 & \cdot \\ \cdot & \cdot & \cdot & \partial_2 \\ \cdot & \cdot & \cdot & 2\partial_1 \end{bmatrix} \begin{bmatrix} \epsilon_{11} \\ \epsilon_{12} + \omega_{12} \\ \epsilon_{12} - \omega_{12} \\ \epsilon_{22} \end{bmatrix} = \boldsymbol{\mathcal{L}}'_\kappa \boldsymbol{v} \tag{42}$$

showing that, for these elements, the operator $\boldsymbol{\mathcal{L}}_\kappa$ is not uniquely defined for given $\boldsymbol{\kappa}$ and $\boldsymbol{v}$.

While the choice of $\boldsymbol{\mathcal{L}}_\kappa$ does not affect the computation of $\boldsymbol{\kappa}$, it does affect the computation of the approximation $\bar{\boldsymbol{\kappa}}$, since

$$\bar{\boldsymbol{\kappa}} = (\boldsymbol{\mathcal{L}}_\kappa \boldsymbol{N}_v)\boldsymbol{v}^N = \begin{bmatrix} \partial_1 \boldsymbol{n}_v & \cdot & \cdot & \cdot \\ \cdot & \partial_2 \boldsymbol{n}_v & \cdot & \cdot \\ \partial_2 \boldsymbol{n}_v & \partial_1 \boldsymbol{n}_v & \cdot & \cdot \\ \cdot & \cdot & \partial_1 \boldsymbol{n}_v & \cdot \\ \cdot & \cdot & \cdot & \partial_2 \boldsymbol{n}_v \\ \cdot & \cdot & \partial_2 \boldsymbol{n}_v & \partial_1 \boldsymbol{n}_v \end{bmatrix} \begin{bmatrix} \boldsymbol{\epsilon}^N_{11} \\ \boldsymbol{\epsilon}^N_{12} + \boldsymbol{\omega}^N_{12} \\ \boldsymbol{\epsilon}^N_{12} - \boldsymbol{\omega}^N_{12} \\ \boldsymbol{\epsilon}^N_{22} \end{bmatrix} = \begin{bmatrix} (\partial_1 \boldsymbol{n}_v)\boldsymbol{\epsilon}^N_{11} \\ (\partial_2 \boldsymbol{n}_v)(\boldsymbol{\epsilon}^N_{12} + \boldsymbol{\omega}^N_{12}) \\ (\partial_2 \boldsymbol{n}_v)\boldsymbol{\epsilon}^N_{11} + (\partial_1 \boldsymbol{n}_v)(\boldsymbol{\epsilon}^N_{12} + \boldsymbol{\omega}^N_{12}) \\ (\partial_1 \boldsymbol{n}_v)(\boldsymbol{\epsilon}^N_{12} - \boldsymbol{\omega}^N_{12}) \\ (\partial_2 \boldsymbol{n}_v)\boldsymbol{\epsilon}^N_{22} \\ (\partial_1 \boldsymbol{n}_v)\boldsymbol{\epsilon}^N_{22} + (\partial_2 \boldsymbol{n}_v)(\boldsymbol{\epsilon}^N_{12} - \boldsymbol{\omega}^N_{12}) \end{bmatrix} \tag{43}$$

$$\bar{\boldsymbol{\kappa}}' = (\boldsymbol{\mathcal{L}}'_\kappa \boldsymbol{N}_v)\boldsymbol{v}^N = \begin{bmatrix} \partial_1 \boldsymbol{n}_v & \cdot & \cdot & \cdot \\ \cdot & \partial_2 \boldsymbol{n}_v & \partial_2 \boldsymbol{n}_v & -\partial_1 \boldsymbol{n}_v \\ 2\partial_2 \boldsymbol{n}_v & \cdot & \cdot & \cdot \\ -\partial_2 \boldsymbol{n}_v & \partial_1 \boldsymbol{n}_v & \partial_1 \boldsymbol{n}_v & \cdot \\ \cdot & \cdot & \cdot & \partial_2 \boldsymbol{n}_v \\ \cdot & \cdot & \cdot & 2\partial_1 \boldsymbol{n}_v \end{bmatrix} \begin{bmatrix} \boldsymbol{\epsilon}^N_{11} \\ \boldsymbol{\epsilon}^N_{12} + \boldsymbol{\omega}^N_{12} \\ \boldsymbol{\epsilon}^N_{12} - \boldsymbol{\omega}^N_{12} \\ \boldsymbol{\epsilon}^N_{22} \end{bmatrix} = \begin{bmatrix} (\partial_1 \boldsymbol{n}_v)\boldsymbol{\epsilon}^N_{11} \\ -(\partial_1 \boldsymbol{n}_v)\boldsymbol{\epsilon}^N_{22} + (\partial_2 \boldsymbol{n}_v)(2\boldsymbol{\epsilon}^N_{12}) \\ 2(\partial_2 \boldsymbol{n}_v)\boldsymbol{\epsilon}^N_{11} \\ -(\partial_2 \boldsymbol{n}_v)\boldsymbol{\epsilon}^N_{11} + (\partial_1 \boldsymbol{n}_v)(2\boldsymbol{\epsilon}^N_{12}) \\ (\partial_2 \boldsymbol{n}_v)\boldsymbol{\epsilon}^N_{22} \\ 2(\partial_1 \boldsymbol{n}_v)\boldsymbol{\epsilon}^N_{22} \end{bmatrix} \tag{44}$$

The approximation $\bar{\boldsymbol{\kappa}}$ depends on the nodal rotations (that is, on the approximated rotation field), which is physically not correct since $\boldsymbol{\kappa}$ does not depend on the rotation field. Therefore both QU34L4 and QU32L2 introduce a non-physical dependence of the mechanical response on the nodal rotations, as they share the same operator $\boldsymbol{\mathcal{L}}_\kappa$.

The alternative approximation $\bar{\boldsymbol{\kappa}}'$ does not depend on the nodal rotations, but its numerical behaviour is very poor, as every element has three additional zero-energy modes that do not disappear in a multi-element mesh. This is because every element only introduces a single equation to link the nodal rotations, so that in most cases the number of nodal rotations to be defined for a mesh is higher than the number of equations involving the nodal rotations.

5 | A NEW MIXED FORMULATION FOR STRAIN-GRADIENT THEORIES

Since the rotation field $\boldsymbol{\omega}$ is not required to define the (generalised) stress-strain behaviour of the material, an improved formulation can be obtained by not interpolating rotations and instead only interpolating strains. While it is possible to do so within a Form-I formulation, it is simpler and more efficient to use a Form-II formulation.

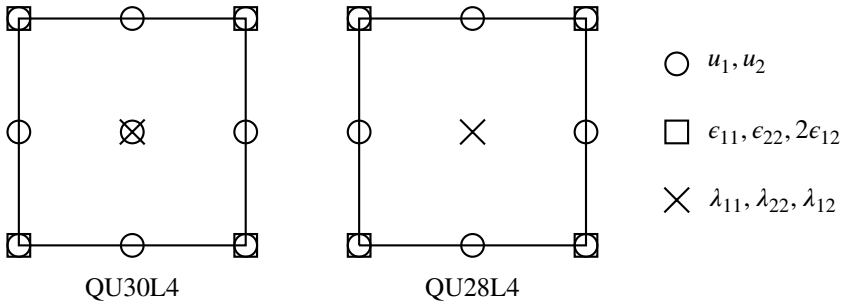

FIGURE 2 Geometry and degrees of freedom for the new QU30L4 and QU28L4 elements.

The relevant continuum quantities can be written as

$$\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \ \boldsymbol{\sigma} = \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{12} \end{bmatrix}, \ \boldsymbol{v} = \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_{11} \\ \epsilon_{22} \\ 2\epsilon_{12} \end{bmatrix}, \ \boldsymbol{\mu} = \begin{bmatrix} \hat{\mu}_{111} \\ \hat{\mu}_{211} \\ \hat{\mu}_{122} \\ \hat{\mu}_{222} \\ \hat{\mu}_{112} \\ \hat{\mu}_{212} \end{bmatrix}, \ \boldsymbol{\kappa} = \begin{bmatrix} \hat{\kappa}_{111} \\ \hat{\kappa}_{211} \\ \hat{\kappa}_{122} \\ \hat{\kappa}_{222} \\ 2\hat{\kappa}_{112} \\ 2\hat{\kappa}_{212} \end{bmatrix}, \ \boldsymbol{\lambda} = \begin{bmatrix} \lambda_{11} \\ \lambda_{22} \\ \lambda_{12} \end{bmatrix} \tag{45}$$

where $\hat{k}_{ijk} = \epsilon_{jk,i}$. The basic differential operators are

$$\boldsymbol{\mathcal{L}}_v = \boldsymbol{\mathcal{L}}_\epsilon = \begin{bmatrix} \partial_1 & \cdot \\ \cdot & \partial_2 \\ \partial_2 & \partial_1 \end{bmatrix}, \ \boldsymbol{\mathcal{L}}_\kappa = \begin{bmatrix} \partial_1 & \cdot & \cdot \\ \partial_2 & \cdot & \cdot \\ \cdot & \partial_1 & \cdot \\ \cdot & \partial_2 & \cdot \\ \cdot & \cdot & \partial_1 \\ \cdot & \cdot & \partial_2 \end{bmatrix} \tag{46}$$

the nodal quantities are

$$\boldsymbol{u}^N = \begin{bmatrix} \boldsymbol{u}_1^N \\ \boldsymbol{u}_2^N \end{bmatrix}, \ \boldsymbol{v}^N = \begin{bmatrix} \boldsymbol{v}_{11}^N \\ \boldsymbol{v}_{22}^N \\ \boldsymbol{v}_{12}^N \end{bmatrix}, \ \boldsymbol{\lambda}^N = \begin{bmatrix} \boldsymbol{\lambda}_{11}^N \\ \boldsymbol{\lambda}_{22}^N \\ \boldsymbol{\lambda}_{12}^N \end{bmatrix} \tag{47}$$

and the shape function matrices are

$$\boldsymbol{N}_u = \begin{bmatrix} \boldsymbol{n}_u & \cdot \\ \cdot & \boldsymbol{n}_u \end{bmatrix}, \ \boldsymbol{N}_v = \begin{bmatrix} \boldsymbol{n}_v & \cdot & \cdot \\ \cdot & \boldsymbol{n}_v & \cdot \\ \cdot & \cdot & \boldsymbol{n}_v \end{bmatrix}, \ \boldsymbol{N}_\lambda = \begin{bmatrix} \boldsymbol{n}_\lambda & \cdot & \cdot \\ \cdot & \boldsymbol{n}_\lambda & \cdot \\ \cdot & \cdot & \boldsymbol{n}_\lambda \end{bmatrix} \tag{48}$$

The rest of the element derivation follows the general procedure presented in section 2. Indeed, a major benefit of the general procedure developed in this paper is that it allows for these new elements to be presented in a very clear and concise way, directly highlighting their similarities and differences with respect to existing elements.

The resulting elements will have fewer degrees of freedom than their counterparts obtained using the existing formulation described in Section 4. Specifically, a QU30L3 and a QU28L3 element are obtained (see figure 2) using the same shape functions as QU34L4 and QU32L4 respectively, but with four fewer dofs (five, including the Lagrange multipliers) in each case. Additionally, both QU30L3 and QU28L3 have no spurious zero-energy modes. Penalty formulations of these elements have also been developed.

The same procedure can be used to obtain a 3D element. With quadratic Lagrangian interpolation, the resulting BR129L6 element will have 129 dofs and 6 Lagrange multipliers, a significant reduction compared to the 153 dofs and 9 Lagrange multipliers of the existing BR153L9 element. A serendipity quadratic interpolation for the displacements would further decrease the number of dofs to 108, leading to a BR108L6 element.

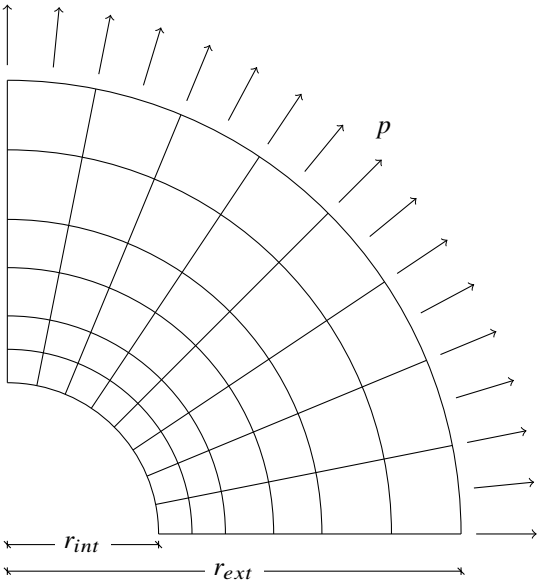

FIGURE 3 FEM model for the thick hollow cylinder benchmark test

6 | NUMERICAL RESULTS

In this section we demonstrate the very good numerical behaviour of the new QU28L3 and QU30L3 elements presented in this paper, and compare it to the behaviour of the existing QU32L4 and QU34L4 elements presented in the literature.

No results are shown for the penalty elements, either existing or newly developed. Initial testing has shown that these elements are able to provide a good approximation to the exact solution of the problem, provided that a good choice is made for the penalty parameter. Such an appropriate choice for the penalty parameter does however depend on element size, so no single value is able to give consistently good results for the wide range of element sizes needed to obtain the convergence plot.

6.1 | Thick hollow cylinder under external traction

All four elements are tested in the benchmark test of a thick hollow cylinder under external distributed loading. Details of the benchmark test, the material model used and the analytical solution, are given by Zervos *et al.*[7]. Quarter symmetry is used, with a typical mesh used shown in figure 3. The internal and external radiuses are $r_{int} = 1$ and $r_{ext} = 3$, the elastic parameters are $E = 1000$ and $\nu = 0.30$, the internal length is $l = 0.5$ and the distributed load is $P_k = p\hat{n}_k$ with $p = 1$ and $\hat{n}_k$ the outward unit normal to the boundary. The ratio of elements in the radial direction (n_r) and in the tangential direction (n_t) is $n_r/n_t = 3/4$. Element size in the radial direction follows a geometric progression with ratio $c = 3^{1/n_r}$. A 3×3 Gauss integration scheme is used for all integrals.

The accuracy of each element type is measured by calculating the relative error between analytical and numerical radial displacement of the outer surface of the hollow cylinder

$$e = \left| \frac{u_{\text{numerical}} - u_{\text{analytical}}}{u_{\text{analytical}}} \right| \tag{49}$$

The analytical solution and the results for increasing mesh density, showing the convergence for each element type, are shown in figure 4. This convergence plot shows that the two new elements (QU28L3 and QU30L3) perform much better than the existing elements (QU32L4 and QU34L4). This is only in small part due to the lower number of degrees of freedom they need for the same mesh. Even considering the number of elements instead of the number of degrees of freedom, the relative error is much smaller for the new elements, a fact therefore attributed to the improved and physically sound formulation of the new elements.

Besides having a smaller relative error, the new elements also show a slightly better convergence rate. Finally, we note that the elements using a serendipity interpolation for the displacements show in general a slightly lower error compared to the elements using a Lagrangian interpolation, while of course also having a lower number of degrees of freedom for the same mesh.

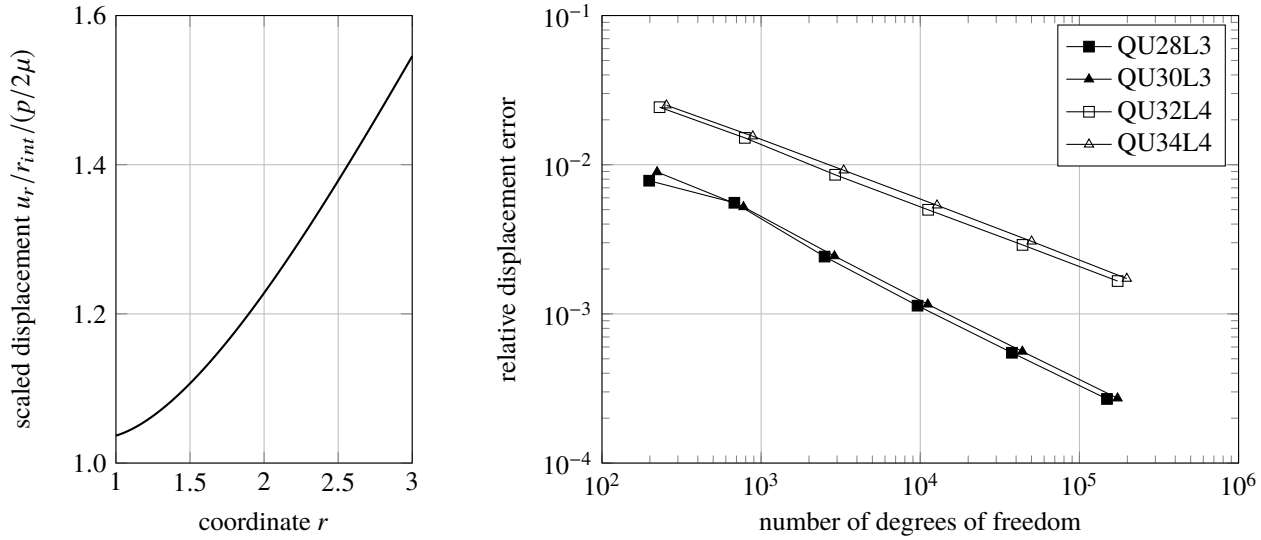

FIGURE 4 Analytical solution and convergence plot for the thick hollow cylinder benchmark test under external traction

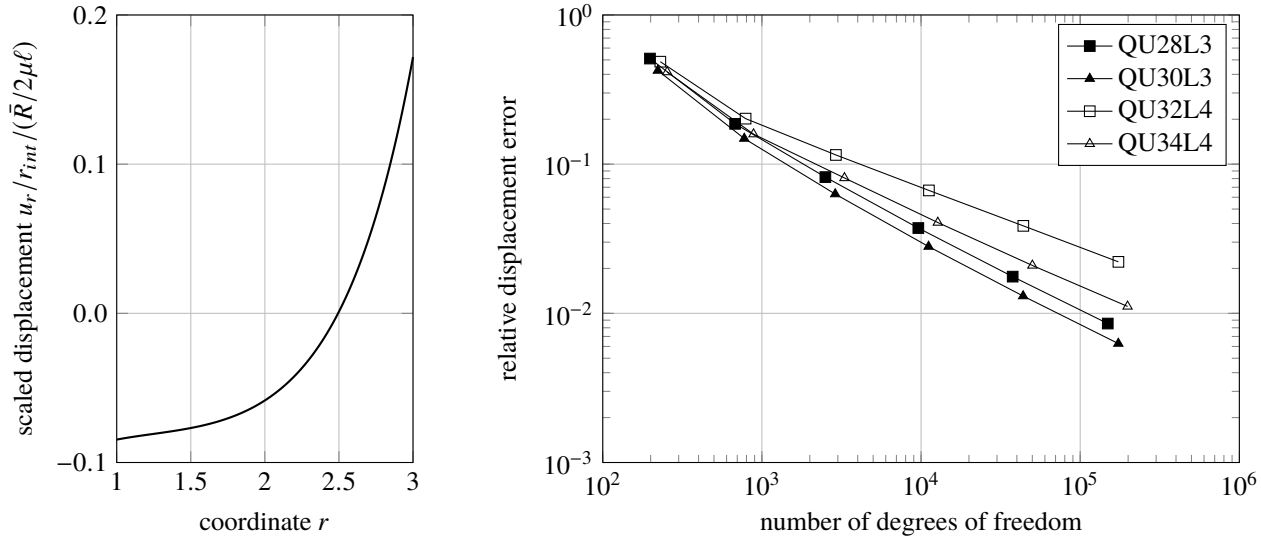

FIGURE 5 Analytical solution and convergence plot for the thick hollow cylinder benchmark test under external double traction

6.2 | Thick hollow cylinder under external double traction

As a second benchmark test, the four elements are tested using the same geometry and meshes as in the first test, but replacing the traction P_i on the external side with a double traction $R_k = \bar{R}\hat{n}_k$ with $\bar{R} = 1$. The analytical solution, obtained from the general case discussed by Papanicolopulos[25], exhibits a more pronounced boundary layer (see figure 5).

The results for increasing mesh density, showing the convergence for each element type, are shown in figure 5. Once more the two new elements (QU28L3 and QU30L3) clearly perform better than the existing elements (QU32L4 and QU34L4), though the difference is smaller than in the previous test. Interestingly, in this case the elements using Lagrange interpolation for the displacements perform better than the ones with serendipity interpolation.

All elements show a higher relative error in this test compared to the previous test. This is probably because the boundary conditions produce work on $\boldsymbol{v}$, instead of $\boldsymbol{u}$, that is on a quantity that is interpolated with a lower-order interpolation and that is only approximately linked with the displacement $\boldsymbol{u}$ that is used to compute the relative error.

7 | CONCLUSIONS

We have presented in this paper for the first time a general framework for developing mixed finite-element formulations for boundary-value problems involving strain-gradient models, using either Lagrange multipliers or penalty methods.

This general framework allows for a unified description of elements previously presented in the literature. It also allows for easy identification of two theoretical issues with the currently proposed elements: spurious zero-energy modes in single elements and dependence on the discretised rotation field. A major benefit of the general framework introduced in this paper is that it makes possible the description of new elements just by presenting a small set of relevant vectors and matrices.

Addressing the issues of existing elements, we present a new family of elements where the strain field is discretised instead of the displacement gradient. These new elements represent a major improvement over all existing ones, as they are simpler, more robust and more efficient. Numerical simulations of a benchmark problem further demonstrate the improved performance of the new elements.

ACKNOWLEDGMENTS

This research was funded from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme under REA grant agreement n^{o} 618096 (Generalised Continuum Models and Plasticity—GECOMPL). F. Gulib also acknowledges the support received from the School of Engineering of the University of Edinburgh.

References

1. Toupin R. A.. Elastic materials with couple stresses. *Archive for Rational Mechanics and Analysis.* 1962;11(1):385-414.
2. Mindlin R. D.. Micro-structure in linear elasticity. *Archive for Rational Mechanics and Analysis.* 1964;16(1):51-78.
3. Mindlin R. D., Eshel N. N.. On first strain-gradient theories in linear elasticity. *International Journal of Solids and Structures.* 1968;4(1):109-124.
4. Froiio F., Zervos A.. Second-grade elasticity revisited. *Mathematics and Mechanics of Solids.* 2018;available online.
5. Shu J. Y., King W. E., Fleck N. A.. Finite elements for materials with strain gradient effects. *International Journal for Numerical Methods in Engineering.* 1999;44(3):373-391.
6. Askes H., Aifantis E. C.. Numerical modelling of size effects with gradient elasticity - Formulation, meshless discretization and examples. *International Journal of Fracture.* 2002;117(4):347-358.
7. Zervos A., Papanicolopulos S.-A., Vardoulakis I.. Two finite element discretizations for gradient elasticity. *Journal of Engineering Mechanics–ASCE.* 2009;135(3):203-213.
8. Papanicolopulos S.-A., Zervos A., Vardoulakis I.. A three dimensional C^1 finite element for gradient elasticity. *International Journal for Numerical Methods in Engineering.* 2009;77(10):1396-1415.
9. Papanicolopulos S.-A.. Chirality in isotropic linear gradient elasticity. *International Journal of Solids and Structures.* 2011;48(5):745-752.
10. Amanatidou E., Aravas N.. Mixed finite element formulations of strain-gradient elasticity problems. *Computer Methods in Applied Mechanics and Engineering.* 2002;191(15–16):1723-1751.
11. Askes H., Gutiérrez M. A.. Implicit gradient elasticity. *International Journal for Numerical Methods in Engineering.* 2006;67(3):400-416.
12. Papanicolopulos S.-A., Zervos A.. Numerical solution of crack problems in gradient elasticity. *Engineering and Computational Mechanics.* 2010;163(2):73-82.

13. Zervos A., Papanastasiou P., Vardoulakis I.. Modelling localisation of deformation and failure with gradient elastoplasticity. In: Cross M., ed. *Proceedings of the 8th Annual Conference of the Association for Computational Mechanics in Engineering (ACME), 16-19 April 2000, University of Greenwich*, ; 2000.

14. Zervos A., Papanastasiou P., Vardoulakis I.. A finite element displacement formulation for gradient elastoplasticity. *International Journal for Numerical Methods in Engineering.* 2001;50(6):1369-1388.

15. Matsushima T., Chambon R., Caillerie D.. Large strain finite element analysis of a local second gradient model: application to localization. *International Journal for Numerical Methods in Engineering.* 2002;54(4):499-521.

16. Kouznetsova V. G., Geers M. G. D., Brekelmans W. A. M.. Multi-scale second-order computational homogenization of multi-phase materials: a nested finite element solution strategy. *Computer Methods in Applied Mechanics and Engineering.* 2004;193(48):5525 - 5550. Advances in Computational Plasticity.

17. Papanicolopulos S.-A., Zervos A., Vardoulakis I.. Discretization of gradient elasticity problems using C^1 finite elements. In: Maugin Gérard A., Metrikine Andrei V., eds. *Mechanics of Generalized Continua*, Advances in Mechanics and Mathematics, vol. 21: New York: Springer 2010 (pp. 269–277).

18. Zybell L., Mühlich U., Kuna M., Zhang Z.L.. A three-dimensional finite element for gradient elasticity based on a mixed-type formulation. *Computational Materials Science.* 2012;52(1):268-273. Proceedings of the 20th International Workshop on Computational Mechanics of Materials – IWCMM 20.

19. Zervos A.. Finite elements for Elasticity with Microstructure and Gradient Elasticity. *International Journal for Numerical Methods in Engineering.* 2008;73(4):564-595.

20. Chakravarty S., Hadjesfandiari A. R., Dargush G. F.. A penalty-based finite element framework for couple stress elasticity. *Finite Elements in Analysis and Design.* 2017;130:65 - 79.

21. Kwon Y.-R., Lee B.-C.. A mixed element based on Lagrange multiplier method for modified couple stress theory. *Computational Mechanics.* 2017;59(1):117–128.

22. Babuška I.. The finite element method with Lagrangian multipliers. *Numerische Mathematik.* 1973;20(3):179–192.

23. Brezzi F.. On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers. *Revue française d'automatique, informatique, recherche opérationnelle. Analyse numérique.* 1974;8(R2):129–151.

24. Zienkiewicz O. C., Taylor R. L.. The finite element patch test revisited: A computer test for convergence, validation and error estimates. *Computer Methods in Applied Mechanics and Engineering.* 1997;149:223-254.

25. Papanicolopulos S.-A.. Analytical and numerical solutions in boundary value problems of materials with microstructure. PhD thesis, National Technical University of Athens, 2008.